

**МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ  
НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ  
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ  
ІМЕНІ ІГОРЯ СІКОРСЬКОГО»**



**ГОРБАТЮК ВЛАДИСЛАВ СЕРГІЙОВИЧ**

УДК 004.032.26

**ІНФОРМАЦІЙНА ТЕХНОЛОГІЯ ПРОГНОЗУВАННЯ  
НЕСТАЦІОНАРНИХ ЧАСОВИХ РЯДІВ НА ОСНОВІ НЕЙРОННИХ  
МЕРЕЖ**

**Спеціальність 05.13.06 – інформаційні технології**

**АВТОРЕФЕРАТ**

**дисертації на здобуття наукового ступеня  
кандидата технічних наук**

Київ – 2021

Дисертацією є рукопис.

Робота виконана на кафедрі технічної кібернетики Національного технічного університету України «Київський політехнічний інститут імені Ігоря Сікорського» Міністерства освіти і науки України.

**Науковий консультант:** доктор технічних наук, доцент  
**Чумаченко Олена Іллівна,**  
Національний технічний університет України  
«Київський політехнічний інститут імені Ігоря Сікорського»,  
доцент кафедри технічної кібернетики

**Офіційні опоненти:** доктор технічних наук, професор  
**Литвиненко Володимир Іванович,**  
Херсонський національний технічний університет,  
завідувач кафедри інформатики і комп'ютерних наук

доктор технічних наук, професор  
**Машталір Сергій Володимирович,**  
Харківський національний університет  
радіоелектроніки, професор кафедри інформатики

Захист відбудеться «16» квітня 2021 р. о 16 годині 00 хвилин на засіданні спеціалізованої вченої ради Д 26.002.29 в Національному технічному університеті України «Київський політехнічний інститут імені Ігоря Сікорського» за адресою: 03056, Київ, пр. Перемоги, 37, корп. №11, ауд. 215.

З дисертацією можна ознайомитися у Науково–технічній бібліотеці ім. Г.І. Денисенка Національного технічного університету України «Київський політехнічний інститут імені Ігоря Сікорського» за адресою: 03056, Київ, пр. Перемоги, 37.

Автореферат розісланий «11» березня 2021 р.

Вчений секретар  
спеціалізованої вченої ради Д 26.002.29  
доктор технічних наук, професор



Теленик С.Ф.

## ЗАГАЛЬНА ХАРАКТЕРИСТИКА РОБОТИ

**Актуальність теми.** Використання технологій на основі штучного інтелекту дозволяє позбавитися окремих припущень та недоліків класичних підходів. На сьогодні найбільш перспективним напрямком інтелектуальних технологій вважаються штучні нейронні мережі, у тому числі завдяки їх універсальності та вражаючим результатам, отриманим у різноманітних сферах їх використання, таких як: аналіз зображень, відео, тексту, мови та інші. Однак, більшість методів прогнозування на основі штучного інтелекту припускають, що прогнозований процес можна описати використовуючи єдину модель, і таке припущення не завжди виконується.

Питанням прогнозування часових рядів займалися вітчизняні та іноземні вчені: Т.В. Алесинська, Є.В. Бодянський, А.Г. Івахненко, Ю.П. Зайченко, О.Г. Руденко, В.С. Степашко, П.І. Бідюк, F.A. Amir, J.L. Elman, R.A. Fisher, T. Hanazawa, G. Hinton, W.S. McCulloch, A. Waibel та інші.

Класичні підходи до розв'язання задачі прогнозування базуються на використанні теорії ймовірності та математичної статистики і припускають, що модель прогнозованого об'єкту відома з точністю до параметрів (лінійна, поліноміальна, експоненціальна та інші), та припускають стаціонарність прогнозованого об'єкту або конкретний вигляд нестаціонарності.

Доцільною і актуальною є задача розробки нових методів, моделей та інформаційної технології прогнозування нестаціонарних часових рядів з використанням нейронних мереж, що дозволить підвищити точність прогнозу за рахунок використання більш загальних припущень щодо прогнозованого процесу та підвищення стійкості моделі до проблеми перенавчання при малій кількості наявних даних.

**Зв'язок роботи з науковими програмами, планами, темами.** Дисертаційне дослідження виконувалось відповідно до планів НДР, програм і договорів, що виконувались в Національному технічному університеті України «Київський політехнічний інститут імені Ігоря Сікорського»: НДР «Методологія побудови інформаційних систем з інтелектуальною підтримкою прийняття рішень» (номер державної реєстрації 0111U002510), НДР «Розробка методів, алгоритмів структурно-параметричного синтезу гібридних нейронних мереж та створення на її основі інтелектуальних систем розв'язання прикладних задач» (номер державної реєстрації 0117U002595).

**Мета і задачі дослідження.** Метою роботи є підвищення точності прогнозу для нестаціонарних часових рядів шляхом створення нової інформаційної технології прогнозування нестаціонарних часових рядів з використанням методів та моделей на основі нейронних мереж.

*Об'єктом дослідження* є процес прогнозування нестаціонарних часових рядів з використанням методів на основі нейронних мереж.

*Предмет дослідження* – моделі, методи та інформаційна технологія, за допомогою яких здійснюється прогнозування нестаціонарних часових рядів.

Основні завдання роботи, які потрібно розв'язати для досягнення мети:

- аналіз проблеми прогнозування нестационарних часових рядів та існуючих методів її вирішення;
- обґрунтування необхідності розроблення нової інформаційної технології;
- визначення класу нестационарних часових рядів, для якого буде розроблятися інформаційна технологія;
- розробка загального підходу до побудови прогнозуючої моделі для визначеного класу нестационарних часових рядів на основі аналізу властивостей часових рядів, що належать до цього класу, та переваг і недоліків існуючих методів;
- розробка нових та вдосконалення існуючих моделей для апроксимації функцій на базі штучних нейронних мереж, враховуючи специфіку визначеного класу нестационарних часових рядів;
- розробка нових методів побудови прогнозуючих моделей для визначеного класу нестационарних часових рядів на основі розробленого загального підходу та моделей;
- розробка інформаційної технології прогнозування нестационарних часових рядів на основі запропонованих методів та моделей;
- експериментальна перевірка розробленої інформаційної технології.

**Методи дослідження** базуються на положеннях:

- теорії штучних нейронних мереж та скінченних алгебраїчних структур, що дозволило розробити нові типи нейронів і архітектури нейронних мереж;
- теорії оптимізації, дискретної математики та лінійної алгебри, що дозволило удосконалити методи навчання штучних нейронних мереж;
- математичного аналізу, що дозволило виявити властивості аналізованих функцій;
- імітаційного та комп'ютерного моделювання, що дозволило підтвердити достовірність отриманих теоретичних результатів;
- математичної статистики, що дозволило оцінити результати експериментів;
- методів програмування та теорії обчислювальної складності алгоритмів, що дозволило реалізувати програмну частину інформаційної технології прогнозування нестационарних часових рядів.

Як засіб розв'язання поставлених завдань використовувалося математичне та комп'ютерне моделювання. У дослідженнях використовувались програмні пакети Tensorflow, Keras та Matlab, а також мова програмування C#.

### **Наукова новизна одержаних результатів.**

1. Уперше запропоновано модель штучного нейрону, яка відрізняється від існуючих тим, що складається з трьох зважених суматорів та нової функції активації, яка залежить від трьох змінних, що дозволяє отримати модель, яка апроксимує певну просту кусково–лінійну функцію і при цьому дозволяє незалежно налаштовувати параметри гіперплощини, яка задає півпростори з різними лінійними функціями. За рахунок цих відмінностей нова модель має покращені апроксимуючі властивості.

2. Уперше розроблено метод розв'язання задачі прогнозування для нестационарних часових рядів, який базується на вирішенні задачі кластеризації та побудові декількох локальних прогнозуючих моделей, що відрізняється використанням вдосконаленого методу кластеризації та спеціального регуляризуючого критерію. Завдяки використанню цього критерію зменшується вплив розмірності вибірки на якість вихідної прогнозуючої моделі та підвищується стійкість локальних моделей до перенавчання.

3. Вдосконалено метод м'якої кластеризації шляхом використання моделі роздільних гіперповерхонь, що дозволяє вирішувати задачу кластеризації для певного класу критеріїв як задачу оптимізації диференційованої функції від фіксованої кількості аргументів. Таким чином кількість параметрів, які необхідно налаштовувати, не залежить від кількості прикладів, що дає можливість обмежити кількість параметрів за необхідності.

4. Вдосконалено багаторядний метод групового урахування аргументів для вирішення задачі прогнозування шляхом застосування алгоритмів зворотного поширення помилки та вибіркового виключення, що дозволило покращити точність прогнозу.

Наукові результати і висновки, отримані у дисертації, підкріплено порівняльними тестами результатів вирішення прикладних задач, отриманих шляхом використання розробленої інформаційної технології та шляхом використання відомих методів та моделей прогнозування нестационарних часових рядів.

**Практичне значення одержаних результатів.** На основі розроблених методів та моделей розроблено спеціалізовану інформаційну технологію для вирішення задач прогнозування нестационарних часових рядів. Інформаційна технологія була апробована шляхом:

1. Використання для прогнозування попиту на лінійку пристроїв релейного захисту МРЗС на державному підприємстві «ВО Київприлад», що дозволило визначити попит на комплектуючі та витратні матеріали і тим самим скоротити витрати (акт впровадження від 12.06.2018).

2. Використання в навчальному процесі кафедри авіаційних комп'ютерно-інтегрованих комплексів факультету аеронавігації, електроніки та телекомунікацій Національного авіаційного університету, зокрема у навчальному курсі «Нейротехнології в комп'ютерно-інтегрованих комплексах» (акт впровадження від 25.09.2020).

**Особистий внесок здобувача.** Основні положення і результати дисертаційної роботи одержані автором самостійно. У друкованих працях, написаних у співавторстві, авторові належить: [1] – аналіз проблеми прогнозування нестационарних часових рядів та існуючих методів її вирішення; [2] – розробка вдосконаленої моделі прогнозування на основі нейронних мереж та застосування методу групового урахування аргументів; [3] – розробка спеціального регуляризуючого критерію, що застосовується для побудови локальних моделей; [4] – розробка вдосконаленого багаторядного алгоритму методу групового урахування аргументів; [5] – розробка загального підходу до

побудови прогнозуючої моделі; [6] – розробка вдосконаленого методу м'якої кластеризації; [7] – розробка нової моделі штучного нейрону sigmoid piecewise; [8] – розробка методу прогнозування, що базується на використанні декількох локальних прогнозуючих моделей; [9] – розробка алгоритму побудови прогнозуючої моделі, який застосовує динамічні ваги, що є варіантом загального підходу використання декількох локальних моделей; [10] – розробка алгоритму прогнозування на базі штучних нейронних мереж; [11] – дослідження існуючих методів прогнозування, що використовують декілька локальних моделей – нейронних мереж; [12] – аналіз переваг та недоліків застосування штучних нейронних мереж у якості моделей прогнозування.

**Апробація результатів дисертації.** Основні результати дисертації доповідались та обговорювались на 14 міжнародних і українських конференціях, зокрема: 4th International Conference on Actual Problems of Unmanned Aerial Vehicles Developments (APUAVD–2017). – Kyiv: NAU, October 17 – 19, 2017; Міжнародна науково–практична конференція «Інформаційні технології та комп'ютерне моделювання». – м. Івано–Франківськ: Прикарпатський національний університет імені Василя Стефаника, 14–19 травня, 2018 р.; XVIII International Scientific Conference «Corporate Governance – Theory and practice». – Kraków: AGH University of Science and Technology, June 21–22, 2018.

**Публікації.** За матеріалами дисертації опубліковано 26 друкованих праць, серед яких 1 монографія, 11 статей (4 з них у закордонних наукових виданнях країн, що входять до ЄС, і є у наукометричній базі даних Scopus, 5 – у фахових наукових виданнях України, що входять до бібліографічної бази даних CiteFactor, Index Copernicus та інших, 2 – у фахових виданнях України з технічних наук, які представлені у базах даних: DOAJ; Ulrich's Periodicals Directory; eLibrary; Google Scholar, ResearchBib, UIF), 14 тез доповідей на наукових конференціях.

**Структура та обсяг дисертації.** Робота складається зі вступу, чотирьох розділів та висновків до них, загальних висновків, що містять основні результати, списку використаних джерел (117 найменувань) та 3 додатків. Загальний обсяг роботи складає 160 сторінок. Основний зміст дисертаційної роботи викладений на 130 сторінках і містить 79 рисунків, 6 таблиць.

## ОСНОВНИЙ ЗМІСТ РОБОТИ

У **вступі** обґрунтовано актуальність теми дисертаційної роботи, сформульовано мету і задачі дослідження, наукову новизну і практичне значення одержаних результатів. Наведено відомості про впровадження результатів роботи, апробацію, особистий внесок здобувача та публікації.

У **першому розділі** розглянуто проблему прогнозування нестационарних часових рядів та наведено актуальні практичні задачі, що потребують її вирішення, таким чином підтверджуючи актуальність даної проблеми. Визначено клас нестационарних часових рядів з декількома потенційними умовними розподілами, для якого буде розроблятися інформаційна технологія. Завдяки визначенню конкретного класу можна спиратись на його властивості при розробці нових методів, моделей та алгоритмів. Розглянуто існуючі методи

прогнозування нестационарних часових рядів і основні їх переваги та недоліки при застосуванні до рядів з визначеного класу.

У роботі розглядається авторегресійна постановка задачі прогнозування, згідно з якою робиться припущення, що майбутні спостереження певним чином залежать від наявних спостережень, та розглянуто новий клас нестационарних часових рядів з декількома потенційними умовними розподілами. Математично ряди цього класу можна описати наступним чином:

1) Існує фіксована кількість потенційних умовних розподілів  $p_i(y_{t+h} / y_t, \dots, y_1), i = 1, \dots, K$ .

2) Кожне значення  $y_{t+h}$  часового ряду генерується згідно одного з потенційних умовних розподілів, який назовемо «активним розподілом в момент часу  $t+h$ »:

$$p(y_{t+h} / y_t, \dots, y_{t-k+1}) = p_{a(t+h)}(y_{t+h} / y_t, \dots, y_{t-k+1}), a(t+h) \in \{1, \dots, K\},$$

де  $a(t+h)$  – функція, що повертає номер активного розподілу в момент часу  $t+h$ .

3. З часом зміна активного розподілу виникає рідко і після виникнення зміни новий активний розподіл не буде змінюватися.

Основні існуючі методи та алгоритми прогнозування, які можна застосовувати для побудови прогнозуючих моделей для класу часових рядів з декількома потенційними умовними розподілами: кореляційно–регресійний аналіз, інтегрована модель авторегресії–ковзного середнього (ІАРКС), штучні нейронні мережі (ШНМ), метод групового урахування аргументів (МГУА) та приховані марківські моделі. У результаті аналізу цих методів та алгоритмів виділено наступні основні недоліки:

1. Використання єдиної моделі незалежно від активного умовного розподілу замість використання оптимальної моделі для кожного умовного розподілу негативно впливає на точність прогнозу такої єдиної моделі – цей недолік притаманний таким методам, як кореляційно–регресійний аналіз, ІАРКС, ШНМ та МГУА.

2. Наявність передумов, при невиконанні яких застосування методу не рекомендується – характерний недолік наступних методів: кореляційно–регресійний аналіз, ІАРКС, приховані марківські моделі.

3. Занадто спрощена модель для описання нестационарних часових рядів з декількома потенційними умовними розподілами, яка не враховує певні потенційні залежності – характерний недолік прихованих марківських моделей.

4. Використання нелінійних моделей в деяких існуючих методах можна розглядати як застосування декількох простіших моделей для різних частин вхідного простору, що дозволяє отримувати точніші моделі для визначеного класу часових рядів. Однак, у моделях, що використовуються в існуючих методах, функція розбиття вхідного простору на частини пов'язана з відповідними локальними моделями – тобто зміна розбиття простору буде призводити до зміни локальних моделей і навпаки, що є небажаним. Цей недолік притаманний таким методам як ШНМ, МГУА.

Описані недоліки основних існуючих методів та алгоритмів побудови

прогнозуючих моделей обґрунтовують необхідність розроблення інформаційної технології для прогнозування нестационарних часових рядів.

У **другому розділі** досліджено властивості класу нестационарних часових рядів з декількома потенційними умовними розподілами з точки зору подальшого розроблення нових методів побудови прогнозуючих моделей для рядів цього класу та розроблено новий загальний підхід побудови прогнозуючої моделі на основі аналізу властивостей часових рядів з цього класу. Крім того, вдосконалено існуючі та розроблено нові методи та моделі, що можна використовувати для побудови локальних прогнозуючих моделей, враховуючи специфіку визначеного класу нестационарних часових рядів.

Проаналізувавши властивості класу нестационарних часових рядів з декількома потенційними розподілами, можна зробити висновок, що для побудови прогнозуючої моделі потрібно:

- 1) оцінити усі потенційні умовні розподіли  $p_i(y_{k+h}/y_k, \dots, y_1), i = 1, \dots, K$ ;
- 2) маючи на вході вектор послідовних значень часового ряду вигляду  $y_{t:t-k+1} = [y_t, \dots, y_{t-k+1}]$  вміти визначати справжній активний розподіл  $p_{a(t+h)}(y_{t+h}/y_{t:t-k+1})$ , згідно з яким буде згенероване майбутнє значення  $y_{t+h}$ .

Володіючи даною інформацією, ідеальна прогнозуюча модель могла б завжди визначати справжній активний розподіл і на його основі обчислювати оптимальне значення прогнозу  $\hat{y}_{t+h}$ . Отже, маємо наступну загальну схему побудови прогнозуючої моделі, що певним чином апроксимує ідеальну:

- 1) Налаштування компоненту, що для заданого вектору  $y_{t:t-k+1}$  визначає номер  $a(t+h)$  відповідного активного розподілу. Компонент повинен певним чином аналізувати отриманий вектор  $y_{t:t-k+1}$  і оцінювати номер найбільш ймовірного активного розподілу  $a(t+h) = a(t) = \dots = a(t-k+1)$ , який згенерував цей вектор спостережень  $y_{t:t-k+1}$ . Цей компонент назовемо компонентом «оцінки активного розподілу».

- 2) Застосувавши компонент оцінки активного розподілу до усіх векторів  $\{y_{t_i:t_i-k+1}\}$  з навчальної вибірки, ми можемо доповнити кожен навчальний приклад  $\langle y_{t_i:t_i-k+1}, y_{t_i+h} \rangle$  відповідним номером оціненого розподілу  $a(t_i+h)$ :  $\langle y_{t_i:t_i-k+1}, y_{t_i+h}, a(t_i+h) \rangle$ . Тоді можливо відібрати усі приклади, що відповідають певному номеру потенційного умовного розподілу  $p_j(y_{k+h}/y_k, \dots, y_1)$ :  $\{\langle y_{t_i:t_i-k+1}, y_{t_i+h}, a(t_i+h) \rangle: a(t_i+h) = j\}$ , та оцінити прогнозуючу модель для цього розподілу. Прогнозуючу модель, що відповідає певному розподілу  $p_j(y_{k+h}/y_k, \dots, y_1)$  назовемо локальною.

Запропоноване вдосконалення багаторядного алгоритму МГУА, яке полягає в подальшому налаштуванні багаточислової поліноміальної мережі, отриманої в результаті виконання базового алгоритму шляхом застосування алгоритму зворотного поширення помилки та алгоритму «вибіркового виключення» (dropout). Подальше налаштування дозволяє досягнути кращого (з точки зору критерію, що мінімізується) вектору параметрів через те, що базовий



алгоритм налаштовує кожен шар мережі окремо, фіксуючи значення параметрів в усіх інших шарах, що в загальному випадку не гарантує оптимальності отриманих параметрів. Вдосконалений алгоритм складається з наступних кроків:

1. Стадія перед-налаштування, яка фактично співпадає з багаторядним алгоритмом методу групового урахування аргументів. У результаті виконання кроків цієї стадії отримуємо поліноміальну нейронну мережу.

2. Стадія до-налаштування, у ході якої виконується подальше налаштування параметрів мережі.

2.1. Згідно алгоритму вибіркового виключення ваги усіх нейронів попередньо помножуються на коефіцієнт  $\frac{1}{p}$ , де  $p$  – ймовірність виключення відповідного нейрону в процесі навчання.

2.2. Для кожної пари <вхідний вектор, вихідне значення> з навчальної вибірки виконуються два так звані «проходи»: прямий прохід – вектор вхідних значень подається на входи мережі і послідовно обчислюються виходи кожного нейрону, поки не буде обчислений вихід останнього нейрону; зворотній прохід – обчислюються значення похідних функції помилки по усіх параметрах усіх нейронів.

2.3. При цьому, згідно з алгоритмом вибіркового виключення, при виконанні проходів для кожної пари <вхідний вектор, вихідне значення> з навчальної вибірки, з вірогідністю  $p$  виключається кожен нейрон мережі – тобто вважається що усі вхідні та вихідні ваги такого нейрону рівні 0, а отже і похідні функції помилки по цим вагам також будуть рівні 0.

2.4. Обчислені значення похідних сумуються по усім прикладам із навчальної вибірки та значення параметрів нейронів моделі оновлюється згідно з алгоритмом зворотного поширення помилки.

2.5. Після зупинки навчання усі ваги мережі множаться на значення коефіцієнту  $p$ .

У результаті отримано звичайну поліноміальну нейронну мережу, а отже її використання для прогнозування на нових даних виконується як звичайно: вектор нових вхідних даних  $\vec{x}$  подається на вхідний шар мережі, після чого поступово обчислюються виходи усіх нейронів шар за шаром, поки не буде обчислений вихід єдиного нейрону в останньому шарі, який і буде значенням прогнозу мережі.

Розроблено новий штучний нейрон Sigmoid Piecewise (SP), що є більш придатним для побудови прогнозуючих моделей нестационарних часових рядів з декількома потенційними умовними розподілами за рахунок використання трьох зважених суматорів, два з яких виконують роль локальних моделей, де кожна модель є активною на окремому регіоні вхідного простору, а третій суматор задає гіперплощину, яка розділяє вхідний простір на два регіони. У такій моделі параметри гіперплощини, що розділяє простір на регіони, та параметри двох локальних моделей, які є активними на відповідних регіонах, є незалежними, що є важливою перевагою над існуючими моделями штучних нейронів. Отже, SP

нейрон має наступну математичну модель:

$$SP(\vec{x}; \vec{w}^{(+)}, \vec{w}^{(-)}, \vec{h}) = \frac{\langle \vec{w}^{(+)}, \vec{x} \rangle}{1 + e^{-k \langle \vec{h}, \vec{x} \rangle}} + \frac{\langle \vec{w}^{(-)}, \vec{x} \rangle}{1 + e^{k \langle \vec{h}, \vec{x} \rangle}}; k > 0.$$

Ця модель є апроксимацією кусково–лінійної функції, яка складається з роздільної гіперплощини  $\vec{h}$ , та 2 лінійних функцій:  $\vec{w}^{(+)} \cdot \vec{x}, \vec{h} \cdot \vec{x} \geq 0$  та  $\vec{w}^{(-)} \cdot \vec{x}, \vec{h} \cdot \vec{x} < 0$  (апроксимація стає як завгодно точною при  $k \rightarrow \infty$ ). Якщо ввести функцію активації SP нейрону  $AF(S_+, S_-, S_h) = \frac{S_+}{1 + e^{-k S_h}} + \frac{S_-}{1 + e^{k S_h}}$ , де  $S_+, S_-, S_h$  – виходи трьох зважених суматорів, то структура SP нейрону буде мати вигляд, наведений на рис. 1.

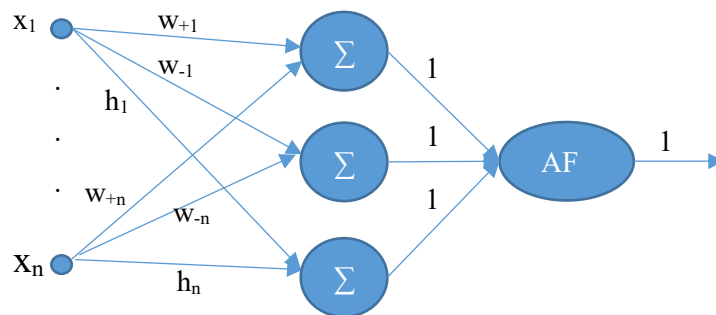


Рис. 1. Структура SP нейрону

Було виконано порівняльні тести ефективності нового нейрону у задачі апроксимації на реальних вибірках по відношенню до ReLU нейрону. Порівняльний тест полягав у навчанні мережі з одним прихованим шаром з відповідними типом нейронів (ReLU або SP), при цьому кількість нейронів у прихованому шарі поступово збільшувалася, поки помилка апроксимації мережі не досягне деякого цільового значення, після чого навчання зупинялося. При кожному збільшенні кількості прихованих нейронів запам'ятовувалася відповідна отримана помилка мережі. Перший порівняльний тест було виконано на часовому ряді денних значень курсу USD до EUR. Графіки залежності кількості нейронів від середньоквадратичної помилки мережі для мережі з ReLU нейронами та мережі з SP нейронами наведені на рис. 2.

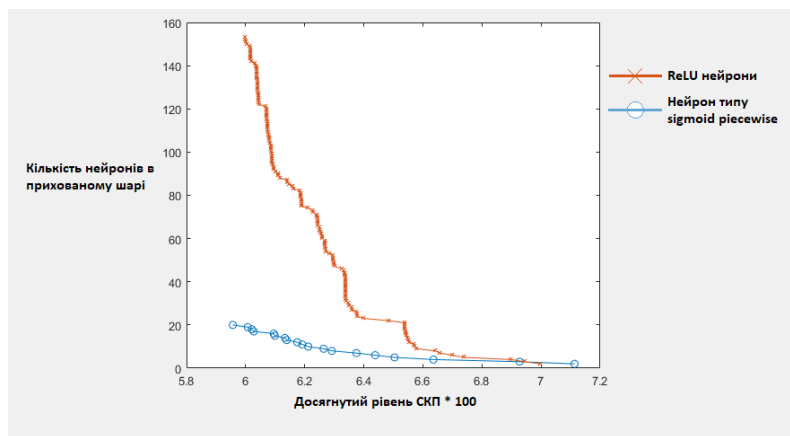


Рис. 2. Залежність кількості нейронів від помилки апроксимації

Як бачимо, темп росту кількості ReLU нейронів у прихованому шарі, потрібних для досягнення певного рівня помилки, значно швидший, ніж відповідний темп для SP нейронів. Другий тест було виконано на вибірці середніх процентних ставок по державним облігаціям в Австралії, 1969–1994 роки. Після виконання цього тесту були отримані наступні графіки, наведені на рис. 3, де видно аналогічну поведінку темпів зростання кількості нейронів:

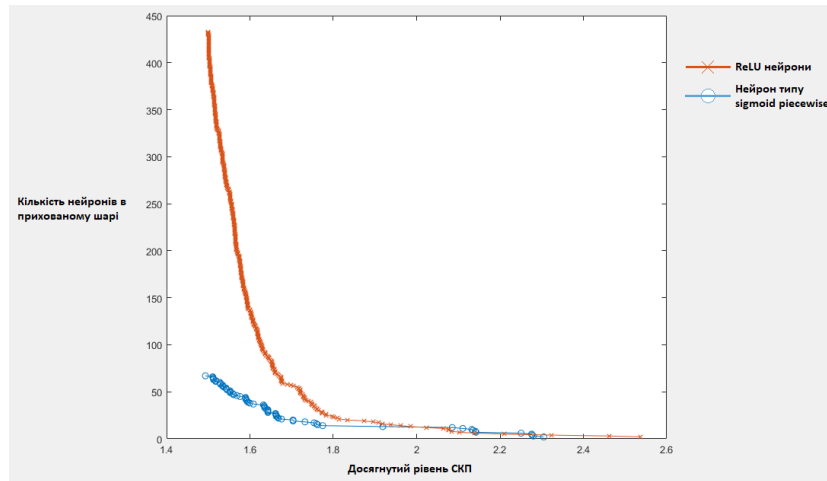


Рис. 3. Залежність кількості нейронів від помилки апроксимації, тест №2

Також було проведене масштабне тестування мереж з SP нейронами на 405 різних часових рядах. У порівнянні брали участь мережа з ReLU нейронами з 12 нейронами в єдиному прихованому шарі та мережа з SP нейронами з 4 нейронами в єдиному прихованому шарі (тобто кількість параметрів для налаштування в обох мережах була однаковою). Для отримання прогнозу використовувалися 4 останні значення часового ряду. Результати наведені у табл. 1 та табл. 2.

Таблиця 1

Результати тестування при прогнозуванні на 5 кроків вперед		
Тип мережі	Мережа з ReLU нейронами	Мережа з SP нейронами
Сумарна помилка мережі на усіх тестових вибірках	121.0895	108.2995

Таблиця 2

Результати тестування при прогнозуванні на 4 кроки вперед		
Тип мережі	Мережа з ReLU нейронами	Мережа з SP нейронами
Сумарна помилка мережі на усіх тестових вибірках	110.2880	99.0999

У **третьому розділі** розглянуто задачу побудови компоненту оцінювача активного розподілу та запропоновано її вирішення шляхом застосування вдосконаленого методу м'якої кластеризації, де вдосконалення полягає у використанні моделі роздільних гіперповерхонь. Крім того, розроблено нові та вдосконалено існуючі методи побудови прогнозуючих моделей для класу

нестационарних часових рядів з декількома потенційними умовними розподілами на основі розробленого загального підходу, методів налаштування локальних моделей та методів побудови компоненту оцінювача активного розподілу:

- новий метод на основі вирішення задачі кластеризації та побудови декількох локальних прогнозуючих моделей, що відрізняється використанням вдосконаленого методу кластеризації та спеціального регуляризуючого критерію;
- вдосконалено метод комбінації експертів (в оригіналі mixture of experts) шляхом використання спеціальної регуляризуючої складової та штучних нейронних мереж з SP нейронами в якості моделей-експертів.

При побудові прогнозуючої моделі згідно розробленої загальної схеми для нестационарних часових рядів з декількома потенційними умовними розподілами потрібно певним чином оцінювати активний умовний розподіл для заданого вектору послідовних значень часового ряду  $y_{t:t-k+1}$ . Одним з типових підходів до вирішення цієї задачі є використання методів та алгоритмів кластеризації для розбиття усієї навчальної вибірки вигляду  $\{ < y_{t_i:t_i-k+1}, y_{t_i+h} > \}, i=1, \dots, N$  на декілька кластерів згідно векторів  $y_{t_i:t_i-k+1}$ , після чого номер активного розподілу для певного вектору буде дорівнювати номеру його кластеру.

Більшість методів кластеризації вирішують задачу дискретної оптимізації певного критерію, і найбільш розповсюдженим є критерій сумарної середньої відстані між точками в одному кластері:

$$CR(\vec{k}, X) = \sum_i^K \sum_{\vec{x}_j: k_j=i} (\vec{x}_j - \vec{\mu}_i)^2,$$

де  $X = (\vec{x}_1, \dots, \vec{x}_n), \vec{x}_i \in R^d$  – множина прикладів,  $\vec{k} = [k_1, \dots, k_n]^T, k_i \in \{1, \dots, n\}$  – вектор номерів кластерів. Цей критерій також можна представити у наступному вигляді:

$$CR(\vec{k}, X) = \sum_i^K \frac{1}{|C_i|} \sum_{\vec{x}_j, \vec{x}_l: k_j=i, k_l=i} (\vec{x}_j - \vec{x}_l)^2.$$

Якщо ввести функцію  $k(\vec{x})$ , що кожному прикладу ставить у відповідність номер його кластеру, то можна також записати цей критерій так:

$$CR(k, X) = \sum_i^K \frac{1}{\sum_{\vec{x} \in X} I(k(\vec{x})=i)} \sum_{\vec{x}_j, \vec{x}_l} \left\{ I(k(\vec{x}_j)=i) \times I(k(\vec{x}_l)=i) \times (\vec{x}_j - \vec{x}_l)^2 \right\},$$

де

$$I(\text{умова}) = \begin{cases} 1, & \text{якщо умова виконується} \\ 0, & \text{якщо умова не виконується} \end{cases}.$$

Розглянемо найпростіший випадок, коли маємо лише 2 кластери, тобто  $k(\vec{x}) \in \{0, 1\}$ . Виконаємо «пом'якшення» вихідної задачі – дозволимо функції  $k(\vec{x})$  приймати значення у всьому діапазоні  $[0, 1]$ , тобто для прикладу  $\vec{x}$  значення функції  $k(\vec{x})$  буде задавати щось подібне до ймовірності приналежності цього

прикладу до першого кластеру; відповідно, значення  $1 - k(\vec{x})$  задає ймовірність приналежності прикладу до другого кластеру. У цьому випадку «пом'якшений» варіант критерію буде мати наступний вигляд:

$$CR(k, X) = \frac{1}{\sum_{\vec{x} \in X} k(\vec{x})} \sum_{\vec{x}_j, \vec{x}_l} \left\{ k(\vec{x}_j) k(\vec{x}_l) (\vec{x}_j - \vec{x}_l)^2 \right\} +$$

$$\frac{1}{\sum_{\vec{x} \in X} [1 - k(\vec{x})]} \sum_{\vec{x}_j, \vec{x}_l} \left\{ [1 - k(\vec{x}_j)] [1 - k(\vec{x}_l)] (\vec{x}_j - \vec{x}_l)^2 \right\}.$$

Якщо ж ми маємо модель роздільної поверхні у вигляді функції  $k(\vec{x}; \vec{w})$ , що залежить від певного вектору параметрів  $\vec{w}$  і є диференційовною по цим параметрам – то критерій  $C(k, X)$  також буде диференційовною функцією від вектору  $\vec{w}$ , а отже для його мінімізації вже можна використовувати апарат мінімізації диференційованих функцій. Отже, пом'якшений варіант задачі кластеризації на 2 кластери можна вирішити як задачу неперервної нелінійної оптимізації, наприклад шляхом використання певної модифікації алгоритму градієнтного спуску. Для розв'язання пом'якшеного варіанту задачі кластеризації на  $K$  кластерів можна застосувати підхід «один проти всіх» – спочатку розділяємо усі приклади на 2 кластери, після чого обираємо кластер з більшою середньою відстанню між його точками, і розбиваємо його на 2 кластери і так далі, поки не отримаємо потрібну кількість кластерів.

Як приклад простої моделі поверхні, що розділяє кластери, можна використати логістичну сигмоїду:  $k(\vec{x}; \vec{w}) = \frac{1}{1 + e^{-\vec{w}^T \vec{x}}}$ . По суті, така модель буде задавати роздільну гіперплощину, що задається вектором параметрів  $\vec{w}$ . Використовуючи цю модель, при мінімізації критерію ми будемо намагатися розділити усі приклади гіперповерхнею на 2 кластери так, щоб сумарна середня відстань між векторами у цих кластерах була мінімальною.

Замість простої моделі логістичної сигмоїди пропонується застосування нейронної мережі з одним прихованим шаром SP нейронів як приклад моделі роздільної гіперповерхні  $k(\vec{x}; \vec{w})$ . Тоді отримуємо наступний опис методу:

1. Виконується розділення усієї множини прикладів на 2 кластери. Для цього:

1.1. Випадковим чином ініціюється початковий вектор параметрів моделі  $\vec{w}_0$ .

1.2. Виконується певна модифікація градієнтного спуску для мінімізації значення критерію  $C(\vec{w})$ , і отримується вектор параметрів  $\vec{w}_f$ .

1.3. Усі приклади розділяються на 2 кластери – ті приклади, для яких значення моделі  $k(\vec{x}; \vec{w}_f) < 0.5$  відбираються у кластер 0, усі інші приклади (тобто такі, для яких  $k(\vec{x}; \vec{w}_f) \geq 0.5$ ) – у кластер 1.

2. Якщо поточна кількість кластерів  $< K$ , то для обох кластерів розраховується їх середня відстань між точками, і кластер з найбільшим

значенням обирається як приклад нової множини прикладів для подальшого розділення на кластери, після чого виконується перехід на перший крок методу. У іншому випадку отримані  $K$  кластерів і є результатом роботи методу.

Пропонується наступний метод побудови прогнозуючих моделей для нестационарних часових рядів з декількома потенційними умовними розподілами на основі вдосконаленого методу кластеризації:

1. Виконується вкладення часового ряду з параметрами  $h$  та  $d$ , у результаті чого отримується матриця вхідних прикладів  $X \in R^{m \times d}$  та вектор цільових змінних  $\vec{o} \in R^m$ , де  $m = n - d - h + 1$ .

2. Навчається мережа з нейронами типу SP в прихованому шарі з використанням всієї множини наявних прикладів. Ваги  $\vec{\theta}^g$  даної мережі є інформацією про глобальний характер прогнозованого процесу.

3. Виконується кластеризація вибірки з використанням вдосконаленого методу м'якої кластеризації. Налаштована модель кластеризації таким чином виконує функцію компоненту оцінювача активного розподілу у загальному підході побудови прогнозуючих моделей для класу нестационарних часових рядів з декількома потенційними умовними розподілами.

4. Навчаються  $K$  нейронних мереж – по одній на кожен кластер. Навчання виконується наступним чином:

а. Ваги кожної мережі ініціалізуються попередньо знайденими вагами  $\vec{\theta}^g$ .

б. Мережа  $net_j$  навчається тільки на прикладах із кластера  $C_j$  з використанням наступної функції помилки:

$$E_j = \sum_{\vec{x}_i \in C_j} (net_j(\vec{x}) - y_i)^2 + \gamma \sum_{\theta_l \in \vec{\theta}_j} (\theta_l - \theta_l^g)^2.$$

Таким чином, при налаштуванні параметрів мережі будуть враховуватися дві складові: помилка мережі на прикладах зі свого кластера і відстань між поточним вектором параметрів мережі та вектором  $\vec{\theta}^g$  – тобто не буде втрачена інформація про глобальний характер процесу. Параметр  $\gamma \in [0, \infty)$  є гіперпараметром, який контролює вплив регуляризуючої складової.

5. Після навчання локальних мереж прогнозування для нових прикладів виконується наступним чином: для вхідного вектора знаходиться відповідний йому кластер  $C_h$  шляхом застосування налаштованої моделі кластеризації, і для прогнозу використовується мережа, навчена на прикладах цього кластера.

Проблемою запропонованого методу побудови прогнозуючих моделей для нестационарних часових рядів з декількома потенційними умовними розподілами на основі використання кластеризації та побудови локальних моделей для кожного кластеру є сильна залежність від якості етапу кластеризації – погана кластеризація може призвести до значного зменшення точності фінальної прогнозуючої моделі. Для зменшення впливу цієї проблеми розроблено модифікацію методу комбінації експертів з використанням аналогічної регуляризації. Модифікований метод комбінації експертів має наступний вигляд:

1. Навчання єдиної глобальної моделі  $f^{(g)}$  із використанням звичайної середньоквадратичної функції помилки. У результаті цього кроку буде отримано вектор параметрів глобальної моделі  $\vec{w}^{(G)}$ .

2. Ініціалізуються  $K$  моделей експертів шляхом «копіювання» глобальної моделі  $f^{(g)}$  – це означає, що у цій модифікації усі експерти повинні мати однакові математичні моделі.

3. Ініціалізується модель компоненту оцінювача активного розподілу  $g$ . Оскільки за своєю функцією ця модель повинна оцінювати важливість усіх моделей експертів для заданого вхідного вектору, де вектор важливостей можна розглядати як дискретний ймовірнісний розподіл, то стає можливим виконати початкове налаштування параметрів стробуючої моделі для виконання м'якої кластеризації навчальної вибірки із застосуванням вдосконаленого методу м'якої кластеризації на основі роздільних гіперповерхонь.

4. Виконується одночасне налаштування параметрів моделей експертів та моделі компоненту оцінювача активного розподілу з використанням наступної функції помилки:

$$RE(\vec{x}, \vec{w}_1^{(e)}, \dots, \vec{w}_K^{(e)}, \vec{w}^{(g)}) = \\ = E(\vec{x}, \vec{w}_1^{(e)}, \dots, \vec{w}_K^{(e)}, \vec{w}^{(g)}) + \gamma * \sum_{j=1}^K \sum_{w_i \in \vec{w}_j^{(g)}} (w_i - w_i^{(G)})^2, \gamma \in R^+.$$

5. Після зупинки стадії налаштування параметрів та отримання фінальної прогнозуючої моделі, прогнозування для нових вхідних векторів виконується шляхом усереднення виходів кожного локального експерту  $f_j$  з вагами, що задаються компонентом оцінювачем активного розподілу  $g$ :

$$f^{(F)}(\vec{x}; \vec{w}_1^{(e)}, \dots, \vec{w}_K^{(e)}, \vec{w}^{(g)}) = \\ = \sum_{j=1}^K g(\vec{x}, j; \vec{w}^{(g)}) f_j(\vec{x}; \vec{w}_j^{(e)}).$$

Для локальних експертів та стробуючої моделі пропонується використання нейронних мереж прямого розповсюдження з SP нейронами у прихованих шарах.

Запропоновані методи були перевірені на штучно згенерованих даних і на реальних вибірках. Для перевірки на штучних даних було згенеровано вибірку наступного вигляду:

$$y_t = \sin(w_t \cdot t) + N(0, \sigma^2), t = 1, \dots, 300,$$

де  $w_{1:100} = \frac{3}{2\pi}, w_{101:200} = \frac{6}{2\pi}, w_{201:300} = \frac{1}{2\pi}, \sigma^2 = 0.01$ , у результаті чого отримано

сигнал, що складається з 3 різних синусоїд з додаванням шуму з нормальним розподілом, графік якого наведено на рис. 4. Очевидно, що цей часовий ряд є неоднорідним. Далі було застосовано метод вкладення часових рядів з розмірністю вкладення  $d = 5$  та прогнозом на 2 кроки вперед, у результаті чого отримано вибірку  $\langle X, y \rangle$ , яку було випадковим чином розбито на навчальну

$\langle X_{train}, y_{train} \rangle$  та тестову  $\langle X_{test}, y_{test} \rangle$ .

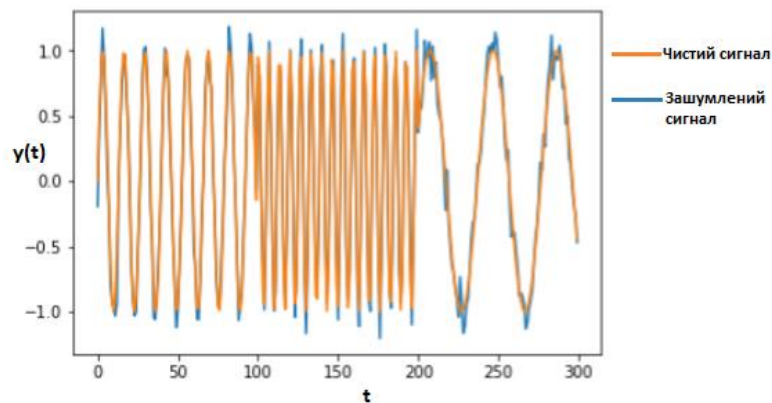


Рис. 4. Згенеровані штучні дані

Після цього було налаштовано 3 різні моделі, помилки яких наведено у наступній таблиці:

Таблиця 3

Помилки моделей на тестовій вибірці на штучних даних

Модель	Середньоквадратична помилка	Середньоабсолютна помилка
Єдина мережа	0.0984	0.2006
Метод комбінації експертів	0.0825	0.1745
Метод регуляризованої комбінації експертів	<b>0.0797</b>	<b>0.1627</b>

Для перевірки на реальних вибірках була використана вибірка з індексом T10Y2Y з масиву даних, що надається онлайн Федеральним Резервним банком Сент-Луїса. Графік вибірки наведено на рис. 5.

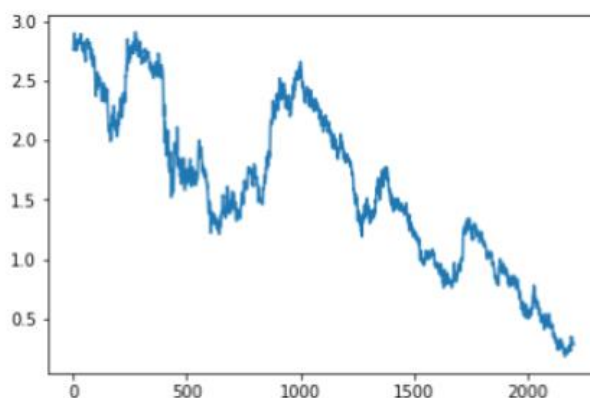


Рис. 5. Вибірка під індексом T10Y2Y, доступ до якої надається онлайн Федеральним Резервним банком Сент-Луїса

Для цієї вибірки було застосовано метод вкладення часових рядів з розмірністю вкладення  $d=10$  та прогнозом на 2 кроки вперед, після чого повторені усі кроки, виконані для штучних даних з тими ж самими значеннями гіперпараметрів. Отримані значення помилок моделей наведені у табл. 4.



Помилки моделей на тестовій вибірці на реальній вибірці

Модель	Середньоквадратична помилка	Середньоабсолютна помилка
Єдина нейронна мережа (двошаровий персептрон)	0.0057	0.0571
Метод комбінації експертів	0.0064	0.0625
Метод регуляризованої комбінації експертів	<b>0.0052</b>	<b>0.0551</b>

Як видно з таблиці, застосування методу комбінації експертів до даного часового ряду призвело до погіршення точності прогнозу моделі по відношенню до застосування єдиної мережі – на практиці ця проблема виникає доволі часто, і вона може бути спричинена багатьма різними факторами, одним з яких є перенавчання локальних експертів. У даному випадку, застосування регуляризованої комбінації експертів дозволило «виправити» цю проблему і отримати найліпшу з точки зору обраних критеріїв модель.

У четвертому розділі розглянуто вимоги та сценарії використання інформаційної технології прогнозування нестаціонарних часових рядів (ПНЧР) на основі штучних нейронних мереж. На основі розглянутих вимог та сценаріїв використання розроблено структурну схему інформаційної технології, схеми реалізації окремих модулів. Модулі ІТ організовано згідно класичної трирівневої архітектури програмних застосунків «Рівень представлення» → «Рівень бізнес-логіки» → «Рівень доступу до даних».

Розроблена ІТ придатна до використання в інформаційних системах з типовою структурою «Користувач» → «Робоча станція» → «Мережа» → «Сервер БД» → «БД». Також можливо, що БД або сховище даних знаходяться на робочій станції, і доступ до них виконується засобами локальної ОС, де сама ІТ застосовується на робочій станції, згідно схеми, наведеної на рис. 6.

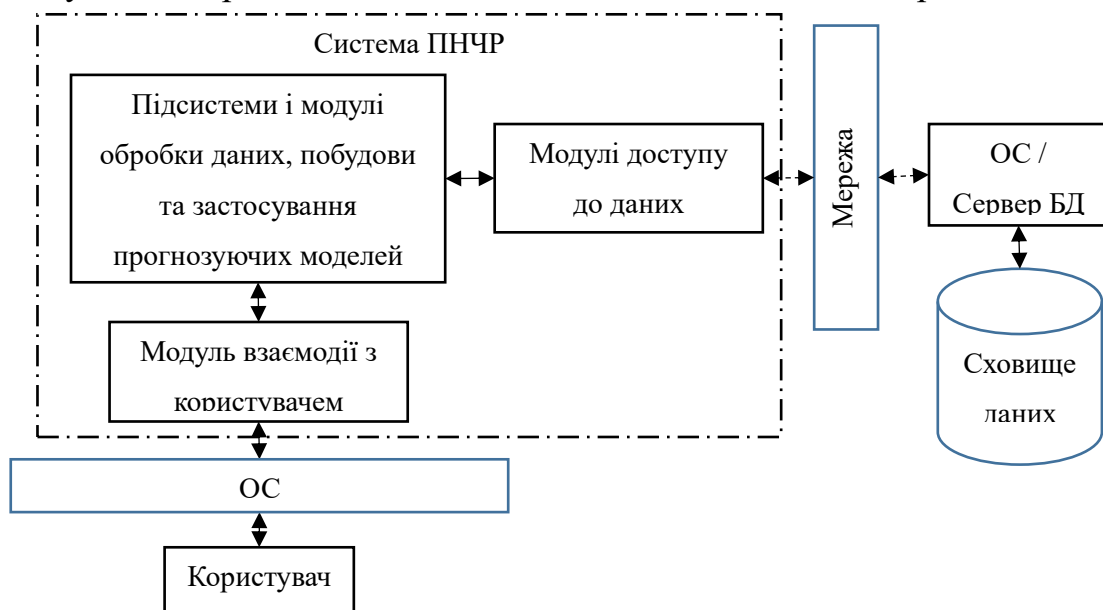


Рис. 6. Інтеграція ІТ ПНЧР в інформаційній системі організації

Для програмної реалізації інформаційної технології ПНЧР на основі використання розроблених методів, алгоритмів та моделей було обрано тип взаємодії «застосунок з графічним інтерфейсом», через такі переваги як зручність використання для недосвідченого користувача та можливість використання складних візуалізацій, які не обмежені можливостями, що надаються веб-браузерами.

До складу системи ПНЧР входять наступні модулі:

- Модуль завантаження даних з зовнішніх джерел: відповідає за відкриття даних із зовнішніх джерел у різноманітних форматах та їх перетворення до внутрішнього формату ІТ.
- Модуль трансформації даних: відповідає за різноманітні засоби обробки та перетворення даних у внутрішньому форматі ІТ.
- Модуль побудови прогнозуючої моделі: відповідає за реалізацію розроблених методів та моделей побудови прогнозуючих моделей для нестационарних часових рядів.
- Модуль застосування прогнозуючої моделі: застосовує побудовану прогнозуючу модель до заданого набору даних.
- Модуль взаємодії з користувачем: реагує на дії користувача та координує роботу інших модулів відповідно до цих дій.
- Модуль візуалізації даних: приймає на вхід певний набір даних та параметри візуалізації та відповідним чином відображає ці дані.
- Модуль збереження даних: перетворює дані з внутрішнього формату до обраного зовнішнього формату та зберігає їх на зовнішньому сховищі.

Цільовою ОС для застосунку обрано Windows XP та новіші версії, для розробки та розгортання застосунку використовується фреймворк Microsoft .NET 4.5. Для підтримки класичних методів побудови прогнозуючих моделей використовуються пакети Cronos та ML.NET, для роботи з ШНМ та реалізації розроблених і вдосконалених у цій роботі методів та моделей – пакет Aforge.NET. Зовнішні дані завантажуються із файлів ОС або з бази даних використовуючи технологію ODBC, що дозволяє використовувати різноманітні типи БД та їх серверів.

ІТ було застосовано для прогнозування попиту на лінійку пристроїв релейного захисту МРЗС на державному підприємстві «ВО Київприлад», що дозволило точніше визначити потребу в комплектуючих та витратних матеріалах і таким чином скоротити витрати. Для побудови моделі було використано наявний часовий ряд щотижневої кількості проданих пристроїв релейного захисту МРЗС за період з 2013 по 2018 рік, графік якого наведено на рис. 7. Для вибору найкращої прогнозуючої моделі було побудовано декілька прогнозуючих моделей використовуючи різноманітні існуючі і нові/вдосконалені методи та алгоритми.

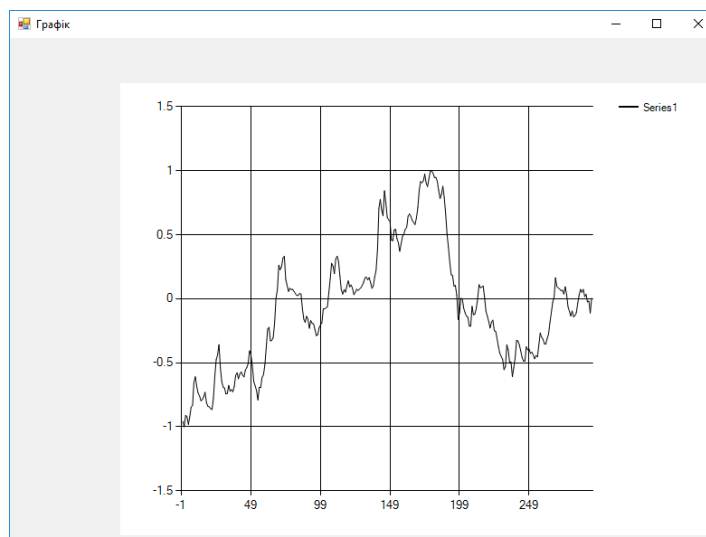


Рис. 7. Нормалізований до діапазону  $[-1,1]$  часовий ряд кількості проданих пристроїв релейного захисту МРЗС

Після налаштування відповідних прогнозуючих моделей були отримані значення їх помилок на тестовій вибірці, які наведені в табл. 5.

Таблиця 5

#### Помилки моделей на тестовій вибірці

Метод/алгоритм, згідно з яким побудовано модель	Середньоквадратична помилка на нормованих даних	Середньоабсолютна помилка на нормованих даних
ARIMA	0.0107	0.1043
Багаторядний алгоритм МГУА	0.0133	0.1157
ReLU нейронна мережа	0.0172	0.1315
SP нейронна мережа	0.0099	0.1011
Кластеризація з побудовою локальних SP нейронних мереж	0.0095	0.0992
Регуляризована комбінація експертів з використанням SP нейронних мереж	<b>0.0087</b>	<b>0.0925</b>

Отже, застосування розробленої інформаційної технології дозволило побудувати прогнозуючу модель, яка на тестовій вибірці мала середньоабсолютну помилку меншу на 11% ніж у найближчого методу – ARIMA.

#### ВИСНОВКИ

Дисертаційна робота становить собою закінчене наукове дослідження, що вирішує актуальну науково–технічну задачу розробки нових методів, моделей та інформаційної технології прогнозування нестационарних часових рядів з використанням нейронних мереж. Основні досягнуті результати:

1. Визначено клас нестационарних часових рядів з декількома потенційними умовними розподілами, досліджено його властивості, а також

переваги і недоліки існуючих методів прогнозування при їх застосуванні до рядів з визначеного класу.

2. Вдосконалено багаторядний алгоритм МГУА шляхом використання методів зворотного поширення помилки та вибіркового виключення, що дає змогу оптимальніше налаштовувати параметри прогнозуючої моделі.

3. Розроблено нову модель штучного нейрону Sigmoid Piecewise (SP), що складається з трьох зважених суматорів та нової функції активації, яка залежить від трьох змінних, що дозволяє отримати модель, яка апроксимує певну просту кусково–лінійну функцію і при цьому дозволяє незалежно налаштовувати параметри гіперплощини, яка задає півпростори з різними лінійними функціями. За рахунок цих відмінностей модель краще підходить для прогнозування рядів з визначеного класу.

4. Вдосконалено метод м'якої кластеризації шляхом використання моделі роздільних гіперповерхонь, що дозволяє вирішувати задачу кластеризації для певного класу критеріїв як задачу оптимізації диференційованої функції. При цьому кількість налаштовуваних параметрів не залежить від кількості прикладів, що дає можливість обмежити кількість параметрів за необхідності.

5. Розроблено новий метод прогнозування рядів з визначеного класу, що використовує вдосконалений метод м'якої кластеризації, SP мережі та спеціальну регуляризацію; вдосконалено метод комбінації експертів шляхом застосування SP мереж та аналогічної регуляризації, що підвищує стійкість експертів до перенавчання.

6. Проведено експериментальні тести розроблених та вдосконалених моделей та методів. Перший тест досліджував ефективність використання SP нейронів для прогнозування реальних вибірок, і показав, що для досягнення певного значення помилки достатньо було використати в 2 рази менше параметрів в SP мережі, ніж в ReLU мережі. У другому тесті досліджувалися помилки прогнозу мереж з SP та ReLU нейронами, і сумарна помилка мереж з SP нейронами на 405 тестових вибірках була на 10% меншою за відповідну сумарну помилку мереж з ReLU нейронами. У третьому тесті порівнювалися новий метод прогнозування на основі кластеризації і побудови локальних моделей, вдосконалений метод комбінації експертів та існуючі методи на штучних та реальних даних. Згідно з результатами тесту, новий та вдосконалений метод досягли зменшення помилки прогнозу на 19% на штучних і на 9% на реальних даних по відношенню до існуючих методів.

7. Створено інформаційну технологію для прогнозування рядів з визначеного класу на основі розроблених і вдосконалених методів та моделей. ІТ було впроваджено для прогнозування попиту на лінійку пристроїв релейного захисту МРЗС на державному підприємстві «ВО Київприлад», що дозволило побудувати прогнозуючу модель з похибкою прогнозу на тестовій вибірці на 11% менше, ніж у найкращого існуючого методу, за рахунок чого точніше оцінено потребу в комплектуючих та витратних матеріалах і скорочено витрати.

## СПИСОК ОПУБЛІКОВАНИХ ПРАЦЬ ЗА ТЕМОЮ ДИСЕРТАЦІЇ

Монографії:

1. Интеллектуальные методы прогнозирования [Текст]: монография / В. М. Синеглазов, Е. И. Чумаченко, В. С. Горбатюк. – К.: Освіта України, 2013. – 219 с. *(Особистий внесок: аналіз проблеми прогнозування нестационарних часових рядів та існуючих методів її вирішення).*

У виданнях іноземних держав–членів ЄС:

2. An algorithm for solving the problem of forecasting / Victor Sineglazov, Elena Chumachenko, Vladyslav Gorbatiuk // Aviation. – 2013. №1(17), p. 9–13. (Видання іноземної держави–члена ЄС, Литва. Входить до міжнародних наукометричних баз ESCI, DOAJ, EBSCOhost, Elsevier Bibliographic Databases: SCOPUS, El Compendex Gale, Google Scholar, ICONDA, Microsoft Academic, MyScienceWork, OpenAIRE2020, ProQuest). *(Особистий внесок: розробка вдосконаленої моделі прогнозування на основі нейронних мереж та методу групового урахування аргументів).*

3. Using a mixture of experts' approach to solve the forecasting task / Victor Sineglazov, Elena Chumachenko, Vladyslav Gorbatiuk // Aviation. – 2014. №3(18), p. 129–133. (Видання іноземної держави–члена ЄС, Литва. Входить до міжнародних наукометричних баз ESCI, DOAJ, EBSCOhost, Elsevier Bibliographic Databases: SCOPUS, El Compendex Gale, Google Scholar, ICONDA, Microsoft Academic, MyScienceWork, OpenAIRE2020, ProQuest). *(Особистий внесок: розробка спеціального регуляризуючого критерію, що застосовується для побудови локальних моделей).*

4. Forecasting aircraft miles flown time series using deep learning-based hybrid approach / Victor Sineglazov, Elena Chumachenko, Vladyslav Gorbatiuk // Aviation. – 2018. № 1(22), p. 6–12. (Видання іноземної держави–члена ЄС, Литва. Входить до міжнародних наукометричних баз ESCI, DOAJ, EBSCOhost, Elsevier Bibliographic Databases: SCOPUS, El Compendex Gale, Google Scholar, ICONDA, Microsoft Academic, MyScienceWork, OpenAIRE2020, ProQuest). *(Особистий внесок: розробка вдосконаленого багаторядного алгоритму методу групового урахування аргументів).*

5. An algorithm for the corporate shares price forecasting / V. Sineglazov, O. Chumachenko, V. Gorbatiuk // Information Technology in Selected Areas of Management 2017. – 2018, p. 73–79. (Видання іноземної держави–члена ЄС, Польща. Входить до міжнародної наукометричної бази Scopus, Google Scholar). *(Особистий внесок: розробка загального підходу до побудови моделі).*

У фахових наукових виданнях України:

6. Soft clustering algorithm based on separating hypersurfaces / O. I. Chumachenko, V. S. Gorbatiuk // Electronics and control systems. – 2017. Volume 2, № 52, p. 11–15. (Наукове фахове видання з технічних наук. Входить до міжнародних наукометричних баз РИНЦ, Google Scholar, EBSCO, Index Copernicus, Наукова періодика України). *(Особистий внесок: розробка вдосконаленого методу м'якої кластеризації).*

7. Structural–parametric synthesis of the feedforward neural networks with sigmoid piecewise–type neurons / M. Z. Zgurovsky, O. I. Chumachenko, V. S. Gorbatiuk // *Electronics and control systems*. – 2018. Volume 4, № 58, p. 42–47. (Наукове фахове видання з технічних наук. Входить до міжнародних наукометричних баз РІНЦ, Google Scholar, EBSCO, Index Copernicus, Наукова періодика України). *(Особистий внесок: розробка нової моделі штучного нейрону sigmoid piecewise)*.

8. Комплексування декількох алгоритмів під час розв’язання задачі прогнозування / О. І. Чумаченко, В. С. Горбатюк // *Житомир: Вісник ЖДТУ. Серія: Технічні науки*. – 2016. №1 (76), С. 101–106. (Наукове фахове видання з технічних наук. Входить до міжнародних наукометричних баз Index Copernicus, DOAJ, Google Scholar, WorldCat, BASE, eLibrary, Наукова періодика України). *(Особистий внесок: розробка методу прогнозування, що базується на використанні декількох локальних прогнозуючих моделей)*.

9. A method for building a forecasting model with dynamic weights / Victor Sineglazov, Elena Chumachenko, Vladyslav Gorbatiuk // *Eastern–European Journal of Enterprise Technologies*. – 2014. [S.l.], v. 2, n. 4(68), p. 4–8. (Наукове фахове видання з технічних наук. Входить до міжнародних наукометричних баз Scopus, CrossRef, Index Copernicus, DOAJ, Google Scholar, OpenAIRE, WorldCat, CORE, SIS, CAS, BASE, EBSCO, Cite Factor, Наукова періодика України). *(Особистий внесок: розробка алгоритму побудови прогнозуючої моделі, який застосовує динамічні ваги)*.

10. Алгоритм решения задачи прогнозирования / Е.И. Чумаченко, В.С. Горбатюк // *Штучний інтелект*. – 2012. № 2, С. 24–31. (Наукове фахове видання з технічних наук. Входить до міжнародних наукометричних баз Google Scholar, Наукова періодика України). *(Особистий внесок: розробка алгоритму прогнозування на базі штучних нейронних мереж)*.

11. Метод решения задачи прогнозирования на основе комплексирования оценок / В. М. Синеглазов, Е. И. Чумаченко, В. С. Горбатюк // *Індуктивне моделювання складних систем*. – 2012. Випуск 4, С. 214–223. (Наукове фахове видання з технічних наук. Входить до міжнародних наукометричних баз РІНЦ, Google Scholar, Наукова періодика України). *(Особистий внесок: дослідження існуючих методів прогнозування, що використовують декілька локальних моделей–нейронних мереж)*.

12. Использование искусственных нейронных сетей для задачи прогнозирования / Е. И. Чумаченко, В. С. Горбатюк // *Електроника и системы управления*. – 2012. № 1, С. 113–119. (Наукове фахове видання з технічних наук. Входить до міжнародних наукометричних баз РІНЦ, Google Scholar, EBSCO, Index Copernicus, Наукова періодика України). *(Особистий внесок: аналіз переваг та недоліків застосування штучних нейронних мереж в якості моделей прогнозування)*.

Матеріали конференцій:

13. One approach for the forecasting task solution / Victor Sineglazov, Elena Chumachenko, Vladyslav Gorbatiuk // *Безпека в авіації та космічні технології: тези*

доповідей V Всесвітнього конгресу «Авіація у XXI столітті». – 2012, р. 3.5.49–3.5.53.

14. Method for predicting failure risk of UAV navigation systems / Elena Chumachenko, Vladyslav Gorbatyuk // Methods and systems of navigation and motion control (MSNMC – 2012). Book of abstracts. – 2012, p. 63–65.

15. Алгоритм решения задачи прогнозирования / Е.И. Чумаченко, В.С. Горбатюк // Интеллектуальные системы принятия решений и проблемы вычислительного интеллекта (ISDMCI – 2012): тезисы докладов. – 2012, С. 428–430.

16. Интеллектуальная система прогнозирования рисков послеоперационных осложнений / В. М. Синеглазов, Е. И. Чумаченко, В. С. Горбатюк // Интеллектуальные системы принятия решений и проблемы вычислительного интеллекта (ISDMCI – 2013): тезисы докладов. – 2013, С. 289–291.

17. Applying Different Neural Network's Topologies to the Forecasting Task / Victor Sineglazov, Elena Chumachenko, Vladyslav Gorbatyuk // International Conference in Inductive Modelling (ICIM – 2013). Book of abstracts. – 2013, p. 217–220.

18. Forecasting the Demand for UAV Using Different Neural Networks Topology / E. I. Chumachenko, V. S. Gorbatiuk // The 2–nd International Conference, Actual Problems of Unmanned Air Vehicles Development Proceedings. – 2013, p. 62–64.

19. Комплексування декількох алгоритмів при вирішенні задачі прогнозування / О. І. Чумаченко, В. С. Горбатюк // Інформаційно–комп'ютерні технології – 2016: тези доповідей VIII Міжнародної науково–технічної конференції. – Житомир: ЖДТУ. – 2016, С. 95–96.

20. Метод побудови прогнозуючої моделі з динамічними параметрами / В. М. Синеглазов, О. І. Чумаченко, В. С. Горбатюк // Інформаційні технології та комп'ютерне моделювання: тези доповідей Міжнародної науково–практичної конференції. – Івано–Франківськ: Супрун В. П. – 2016, С. 23–26.

21. Object's movement prediction in 3D space using neural networks / O. Chumachenko, V. Gorbatiuk // 4th International Conference Methods and systems of navigation and motion control (MSNMC–2016). Book of abstracts. – 2016, p. 255–258.

22. A New Approach in Cluster Analysis / Victor Sineglazov, Elena Chumachenko, Vladyslav Gorbatyuk // 4th International Conference on Actual Problems of Unmanned Aerial Vehicles Developments (APUAVD–2017). Book of abstracts. – 2017, p. 223–227.

23. Нова модель штучного нейрону для побудови нейронних мереж у задачі апроксимації / О.І. Чумаченко, В.С. Горбатюк // Інформаційні технології та комп'ютерне моделювання: тези доповідей Міжнародної науково–практичної конференції. – Івано–Франківськ. – 2018, С. 296–299.

24. Алгоритм м'якої кластеризації на основі розділяючих гіперповерхонь / В.С. Горбатюк // Електроніка – 2018: тези доповідей XI Міжнародної науково-технічної конференції молодих вчених. – 2018, С. 425–430.

25. An algorithm for the corporate shares price forecasting / V. Sineglazov, O. Chumachenko, V. Gorbatiuk // XVIII International Scientific Conference «Corporate Governance – Theory and practice». Book of abstracts. – 2018, p. 101–102.

26. Прогнозування часових рядів з використанням регуляризованої суміші експертів / В. М. Синеглазов, О. І. Чумаченко, В. С. Горбатюк // Інформаційні системи та технології ICT–2018: тези доповідей VII Міжнародної науково-технічної конференції. – 2018, р. 457–459.

## АНОТАЦІЯ

**Горбатюк В.С. Інформаційна технологія прогнозування нестаціонарних часових рядів на основі нейронних мереж.** – На правах рукопису.

*Дисертація на здобуття наукового ступеня кандидата технічних наук за спеціальністю 05.13.06 – інформаційні технології. – Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського» МОН України, м. Київ, 2021.*

Робота присвячена розробці нових методів вирішення актуальної науково-прикладної задачі прогнозування часових рядів.

Запропоновано новий загальний метод побудови прогнозуючих моделей для визначеного класу часових рядів. Розроблено нову модель штучного нейрону Sigmoid Piecewise, що підходить для опису визначеного класу часових рядів. Вдосконалено багаторядний алгоритм МГУА шляхом використання методів зворотного поширення помилки та вибіркового виключення. Запропоновано новий алгоритм м'якої кластеризації на основі моделей роздільних гіперповерхонь. На його основі розроблено метод побудови прогнозуючих моделей для визначеного класу часових рядів.

Розроблено інформаційну технологію для прогнозування нестаціонарних часових рядів, яку було впроваджено для задачі прогнозування попиту на лінійку пристроїв релейного захисту МРЗС на державному підприємстві «ВО Київприлад».

**Ключові слова:** штучні нейронні мережі, прогнозування нестаціонарних часових рядів, метод групового урахування аргументів, інформаційна технологія, метод комбінації експертів, м'яка кластеризація, нелінійна оптимізація.

## АННОТАЦИЯ

**Горбатюк В.С. Информационная технология прогнозирования нестаационарных временных рядов на основе нейронных сетей.** – На правах рукописи.

*Диссертация на соискание ученой степени кандидата технических наук по специальности 05.13.06 – информационные технологии. – Национальный технический университет Украины «Киевский политехнический институт*



*имени Игоря Сикорского», МОН Украины, г. Киев, 2021.*

Работа посвящена разработке новых методов решения актуальной научно–прикладной задачи прогнозирования временных рядов.

Разработана новая модель искусственного нейрона Sigmoid Piecewise, который лучше подходит для описания выделенного класса временных рядов. Усовершенствован многорядный алгоритм МГУА путем использования методов обратного распространения ошибки и выборочного исключения. Предложен новый алгоритм мягкой кластеризации на основе моделей отдельных гиперповерхностей.

На основе созданных методов и моделей разработана информационная технология для прогнозирования нестационарных временных рядов, которая была использована для задачи прогнозирования спроса на линейку устройств релейной защиты МРЗС на государственном предприятии «ПО Киевприбор».

**Ключевые слова:** искусственные нейронные сети, прогнозирование нестационарных временных рядов, метод группового учета аргументов, информационная технология, метод комбинации экспертов, мягкая кластеризация, нелинейная оптимизация.

## ABSTRACT

**Horbatiuk V.S. Information technology for non–stationary time series forecasting based on neural networks.** – Manuscript.

*Thesis for a Candidate Degree in Engineering, specialty 05.13.06 – information technologies. – National Technical University of Ukraine «Igor Sikorsky Kyiv Polytechnic Institute» of Ministry of Education and Science of Ukraine, Kyiv, 2021.*

The thesis is devoted to the development of information technology together with new and improved methods and models for the relevant scientific and applied problem of time series forecasting.

Classical approaches to the forecasting problem are based on the use of probability theory and mathematical statistics assuming that the forecasted object's model (linear, polynomial, exponential and others) are precisely known. They also typically assume the predicted object to be stationary or expect a specific type of “nonstationarity”. It is expedient and relevant to develop new methods, models and information technology for non–stationary time series forecasting with the use of neural networks, which will improve the quality of the forecast by having more flexible assumptions about forecasted object and by reducing the influence of overfitting when amount of available data points is limited.

A new class of nonstationary time series is described. Due to focusing on a specific class of time series, its properties can be used when developing new methods, models and algorithms for building forecasting models for the selected time series class. Existing methods of nonstationary time series forecasting have been analyzed, together with their main advantages and disadvantages when applying to the selected class of time series.

A new generic approach for the forecasting models building for time series from the selected class has been developed. The approach is based on the analysis of the

class properties, along with the advantages and disadvantages of existing methods. The approach consists of two high level steps: first, building the model that determines the currently active conditional probability distribution for the given input vector and the second step – building multiple local models based on the previously estimated information of active conditional probability distributions for each sample vector.

The multi-row GMDH algorithm has been improved by using the backpropagation and dropout, which allowed to fine-tune the forecasting model initial parameters thus improving the forecasting accuracy.

A new model of an artificial neuron – Sigmoid Piecewise (SP) neuron has been developed, which is more efficient for time series from the selected class description. It consists of three weighted sum blocks and a new activation function of three variables, which allows to obtain a model that approximates a simple piecewise linear function and to independently adjust parameters, that define the piecewise function's separating hyperplane, and parameters, that define linear functions for the inputs from different sides of the hyperplane. Due to these differences, the new model can better approximate certain functions and is more suitable for time series from the selected class. Experimental tests were carried out on real world time series, in which it was enough to use 2 times fewer parameters in the SP network than in the ReLU network to achieve the same error value. In another test, the error of SP network on the test set was the smallest among all other methods being compared, and 8% less than the error of the second most accurate method (ARIMA).

The soft clustering method has been improved by using the separating hypersurfaces model, which allowed solving the clustering problem for a certain class of criteria as a differentiable function's optimization problem. In this case, the number of parameters that need to be tuned does not depend on the number of samples, making it possible to limit the number of parameters if necessary.

A new method for nonstationary time series forecasting has been developed. The method is based on clustering implementation followed by a construction of multiple local forecasting models. The method uses improved soft clustering method based on separating hypersurfaces model and the special regularization. The added special regularization reduces the influence of the dataset size on the quality of the obtained forecasting model and increases resilience of local models to the overfitting. New and improved methods have been compared with existing ones on the artificial and real datasets.

Based on the created methods and models, the information technology for non-stationary time series forecasting has been developed to forecast the demand for a line of relay protection MRZS devices at the state enterprise «PA Kyivprylad». As a result, the forecasting model had 11% lower mean absolute error than the best corresponding error among tested existing methods. It became possible to more accurately determine the demand for components and consumables, and thereby reduce costs.

**Keywords:** artificial neural networks, non-stationary time series forecasting, group method of data handling, information technology, mixture of experts method, soft clustering, nonlinear optimization.